A STOCHASTIC MODEL ON SUPERCOILING-DEPENDENT TRANSCRIPTION

by
Yuncong Geng

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Master's of Science and Engineering

Baltimore, Maryland
May, 2020

© 2020 Yuncong Geng
All rights reserved

Abstract

Transcription in *E. coli* introduces supercoiling to DNA, while the local supercoiling profile influences transcription activities in turn. This interplay gives rise to emergent properties, for example, the collective behaviors of RNAPs in highly-transcribed genes, the regulation of transcriptional noise by topological domains, and the correlation of transcription between neighboring genes. Here we propose a spatially-resolved stochastic model for transcription, featured with an explicit description of the interaction of RNAP and DNA supercoiling in all stages of transcription, the stochastic domain formation, and the diffusion of supercoiling within the domain. Our model reproduces the recently discovered cooperative behaviors of a group of RNAPs mediated by "supercoiling cancellation" and provides support for the idea that the topological domain serves as a transcriptional regulator. We also look at how the transcription-induced supercoiling shapes the coordination of neighboring genes and how it depends on the relative orientation of genes. All these findings provide insights into the role of promoter architecture and genome organization in regulating gene expression. However, our model is limited in quantitatively recapitulating several experimental observations (like the antagonistic behavior of a group of RNAPs), reflecting some defects in the model assumptions. More considerations need to be taken into the model, which should be aided with further experimental characterizations of mechanical properties of RNAPs and DNA torsional stress.

Thesis Readers

Dr. Elijah Roberts (Primary Advisor)
Assistant Professor
Thomas C. Jenkins Department of Biophysics
Johns Hopkins University

Dr. Jie Xiao
Professor
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University

Dr. Taekjip Ha
Professor
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University

Contents

List of Figures

Chapter 1

Introduction

In the original version of the central dogma, genetic information was considered to be stored purely in the form of DNA sequences and to be transmitted unidirectionally from DNA to RNA to protein. However, recent studies show that the torsional stress in DNA can also store information and thus mediates the "interaction" between DNA and RNA. It has been shown that in *E. coli*, the torsional stress in DNA (also called supercoiling), resulting from overwinding or underwinding of the DNA double helix, regulates transcription [1] and is regulated by transcription in turn [2, 3]. To be specific, during transcription, due to the large resistance to rotation, RNA polymerase (RNAP) keeps overwinding the DNA ahead of it and underwinding the DNA behind it, and local torsional stress is built as a result. If the stress cannot be relieved timely by other cellular processes, transcription machinery will react to those changes and alter the initiation, elongation, and termination behavior. Previous research shows that an underwound promoter will boost transcription initiation [4, 5], while an elongating RNAP might slow down [5], stall [6], or even dissociate from DNA [7] under a high local DNA torsional stress.

As a feedback mediator, DNA supercoiling complicates the information transmission and brings new questions on how transcription is regulated in *E. coli*. Here we mainly consider the role of DNA supercoiling in three levels of gene transcription, on the level

of a single gene, multiple genes, and topological domains, respectively (**Fig. 1-1**):

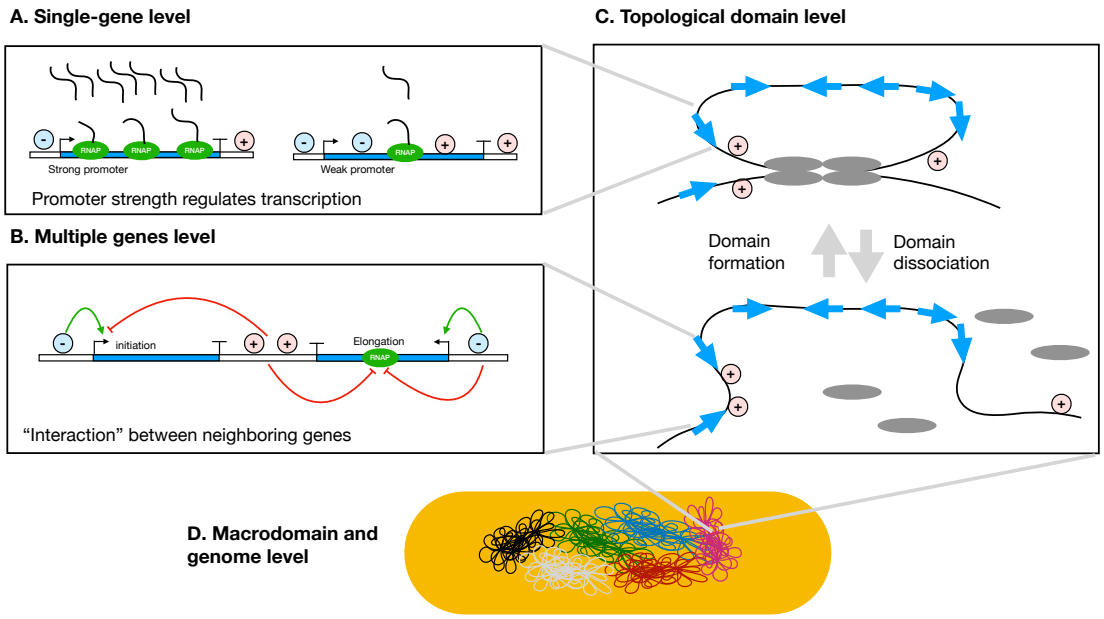

Figure 1-1. Example of transcription regulation at different scales. (A) On the single-gene scale, a gene with a stronger promoter initiates more frequently, elongates more processively, and has more mRNA products. (B) On the multiple genes scale, two convergent genes regulate the transcription of each other: positive supercoils induced by transcription accumulate between the intergenic region; the positive supercoiling temporarily inhibits the initiation of the left gene while the large torque inhibits the elongation of the right gene. (C) On the topological domain scale, the existence of the topological barrier blocks the interactions between two genes and the diffusion of supercoiling. When the domain barrier dissociates, the interaction and diffusion resume.

1) single-gene level:

One of the most important transcription regulators for a gene is its promoter strength. Promoter strength is typically considered to regulate transcription only at the initiation level. However, recent studies show that this regulation also happens at the elongation level: stronger promoters confer an advantage to the processivity of RNAPs when the promoter is on and confer a disadvantage when the promoter is off (Kim *et al.* [7]). This phenomenon is explained by the collective behavior of RNAPs mediated by the dynamics of supercoiling between them. However, it is not yet clear how much supercoiling dynamics can quantitatively explain the experimental results and how

those behaviors depend on the physical properties of DNA and RNAP (for example, the diffusion coefficient of supercoiling and the rotation rate of RNAP).

2) multiple genes level:

In addition to the interaction between co-transcribing RNAPs on a single gene, linearly arranged genes could also interact with each other, similarly, mediated by the intergenic supercoiling dyanmics. Genome-wide evidence of spatial correlation of transcriptional activity has been detected [8, 9], and the transcription of gene cassettes inserted in the *E. coli* genome manifests a strong genome-position dependence [10, 11]. On a smaller scale, it is shown that two closely located divergent genes can increase the transcription of one another since the negative supercoiling generated between them promotes the initiation [9]. Will we observe some collective behaviors between multiple genes? How those behaviors depend on the length, orientation, and intergenic space between the genes? A systematic investigation is needed.

3) topological domain level:

Transcription can also be regulated on a larger scale. It has long been noticed that the 3D genome organization plays a crucial role in regulating gene expression and driving cell-fate decisions in eukaryotic cells [12, 13, 14]. Although the prokaryotic DNA is relatively loosely structured, we hypothesize that the same thing applies to prokaryotic cells. In *E. coli* cells, there are about 400 topological domains (with the average size of 10kb) [15] formed by the dynamic looping of nucleoid-associated proteins. Those proteins become hurdles for the diffusion of supercoiling and make torsional stress constrained to the topological domain. Considering that supercoiling and transcription are closely coupled, we want to know what is the role of topological domains in regulating gene expressions, how it depends on the kinetic rate of domain formation and dissociation and the organization of the genes within the domain.

Addressing the above questions will help us understand the evolutionary design principles for genome organization and also help us design better synthetic circuits by utilizing the mechanical properties of DNA [16]. However, considering the multiple interdependent processes involved in the regulation of transcription and supercoiling, we cannot address the above questions thoroughly without a quantitative framework. Although several models have been proposed [17, 18, 19, 20], none of them fully captured the complicated interactions between transcription and supercoiling.

Here we present a spatially resolved, chemical master-equation based model of *E. coli* transcription, integrating the most updated knowledge about transcription and supercoiling revealed by experiments. This model contains all the processes of transcription (namely initiation, elongation, and termination), their interplay with DNA supercoiling, and other cellular processes that regulate DNA torsional stress (including the topoisomerase activity, supercoiling diffusion, and the dynamical formation and dissociation of topological domains). (**Fig. 1-2**)

The features of our model are as follows. First, we incorporate some crucial results from recent single-molecule studies into the model, like the stall and rotation of RNAP during transcription, which enables us to quantify the transcription apparatus more accurately. Second, it is the first time that we take into consideration the role of dynamic topological domains, which is an important part of the genome organization. Third, we calibrate our model carefully with real-world experimental results.

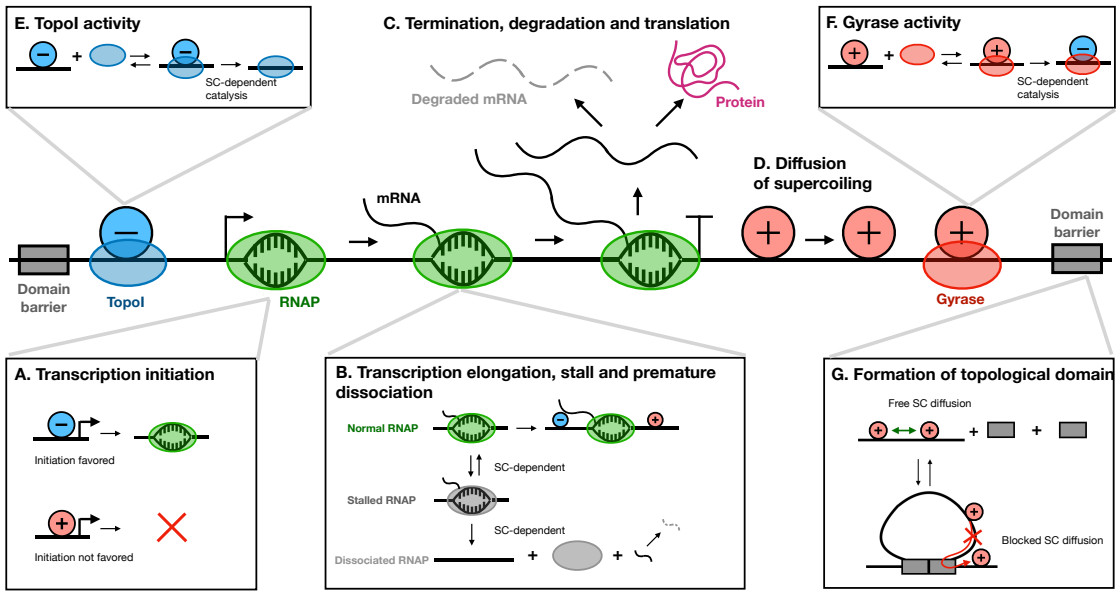

Figure 1-2. Schematic of supercoiling-dependent transcription in a topological domain. Transcription of a gene can be decomposed into three steps: Initiation (A), elongation (B), and termination, degradation and translation (C). Both Initiation and elongation are supercoiling-sensitive. During elongation, the RNAP can switch between normal and stalled state; stalled RNAP has a chance to dissociate from DNA; translocation of RNAP introduces supercoiling. The transcription-induced supercoils can diffuse (D) along DNA and can be removed by TopoI (E) or Gyrase (F). Supercoiling diffusion is confined within the topological domain (G).

Chapter 2

Methods

2.1 General design of the model

In our model, we divided the DNA to 60-bp segments, and we used individual species to represent the molecules at different segments. For example, to model a 3-kb long DNA, we use 50 species (i.e., DNA(1) up to DNA(50)) to represent the availability of DNA on different segments. RNAP(2) represents the presence of RNAP on the second segment, and Turn(3) represents the number of twists and writhes on the third DNA segment. Using this strategy, we can model the translocation of RNAP on the DNA, monitor the changes in supercoiling at different segments, so on and so forth. The detailed rationale of this strategy can be found in **Appendix II**. Reactions were represented in master equations, and simulations were performed with exact sampling (Gillespie algorithm) in LMES [21]. Species counts were recorded every 1 s. Due to the limited number of species allowed in LMES, the maximum size of DNA we can simulate is about 20 kb.

Reaction schemes were built by carefully examining the following two questions: 1) how DNA supercoiling is regulated by transcription and other cellular activities, and 2) how transcription machinery responds to torsional stress. Model assumptions on these two questions will be introduced in the following sections, and all model equations

and parameters can be found in **Appendix I**.

2.2 Characterizing the regulation of DNA supercoiling

As far as we know, three processes are contributing to the dynamics of DNA supercoiling: 1) diffusion of supercoiling, 2) topoisomerase activities, and 3) transcription. Topoisomerase regulates supercoiling globally, transcription perturbs supercoiling locally, and through diffusion, those changes are transmitted along DNA.

1) diffusion of supercoiling:

We use "supercoiling density" (σ) to quantify the level of torsional stress: relaxed DNA makes a turn every 10.5 bp, and supercoiling density is defined as the number of turns added or removed relative to the total number of turns in the relaxed DNA. If the DNA is underwound relative to the relaxed state, the supercoiling density is negative; if DNA is overwound, the supercoiling density is positive. Under a stretching force (which can be provided by an elongating RNAP [22]), a sufficient level of supercoiling density will trigger a phase transition in DNA topology [23]: instead of twisting more densely (or sparsely), DNA will wrap around itself to form writhes (a process called DNA buckling); this is because lower free energy will be achieved when DNA is buckled (**Fig. 2-1 (A)**). Further changes in supercoiling density will cause the writhes to pile up and form plectonemes. The critical supercoiling density for buckling transition is dependent on the stretching force. (**Fig. 2-1 (B)**)

According to the *in vitro* experiment performed by Loenhout *et al.* [24], plectonemes have a very low diffusion coefficient (D), and the diffusion coefficient is force-sensitive (**Fig. 2-1 (C)**): when the DNA is subject to the applied force of 0.5 pN, D is around 0.1 $\mu m^2 \cdot s^{-1}$, and a force up to 2 pN will further reduce D to 0.01 $\mu m^2 \cdot s^{-1}$. Although

no experiments have been done for twists, it is estimated that twists diffuse very fast, with D at about $50 \sim 180\ \mu m^2 \cdot s^{-1}$ [25, 26].

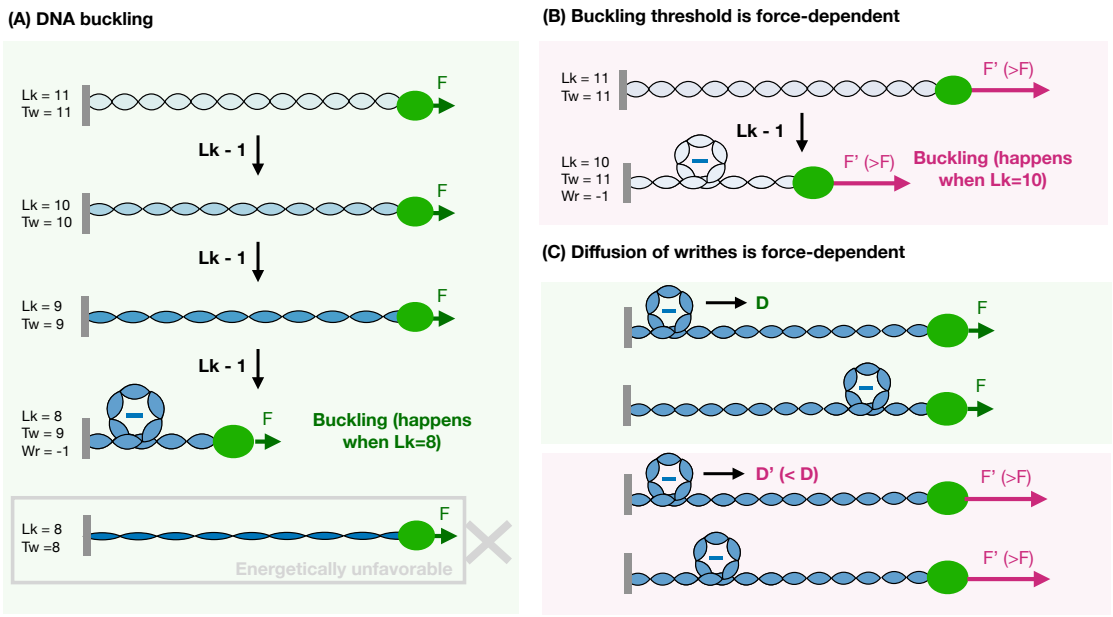

Figure 2-1. Schematic of torsional DNA buckling and diffusion of torsional stress. In (A), the left end of DNA is fixed on a topological barrier, and the right end is pulled by force F. Initially, DNA exists in the form of pure twists. We reduce the linking number by one at each step (it could be removed by topoisomerase or other mechanisms, here we just assume that it happens). At first, DNA reduces the density of twists. When the linking number is reduced from 9 to 8, the DNA buckles itself to form a negative supercoil and preserves the number of twists in the previous step. Compared to (A), if we raise the stretching force to F' at the right end (B), buckling will happen earlier (when the linking number is reduced to 10). In (C), when the stretching force is low (F), writhes diffuse fast on the DNA; when the stretching force is higher (F'), writhes diffuse slower on the DNA.

Here we assume that under physiological conditions, DNA exists mostly in the form of writhes (or plectonemes), and thus torsional stress diffuses at a relatively low speed. For simplicity, we did not distinguish between twists and writhes in our model; rather, we used the universal "turns" to represent them. To reduce the computational load, we modeled a biased random walk, which means that we only allow turns to move down the gradients:

If Turn(k)>Turn(k-1) and DNA(k-1)=1:

$$\text{Turn(k)} \rightarrow \text{Turn(k-1)}$$

If Turn(k)>Turn(k+1) and DNA(k+1)=1:

$$\text{Turn(k)} \rightarrow \text{Turn(k+1)}$$

The diffusivity of turns is determined by the propensity of the above two reactions. We call the rate constant "drift rate", termed as k_{drift}. Considering the molecular crowdedness *in vivo*, we set k_{drift} as 50 s^{-1}, which yields a diffusion coefficient D of 0.02 $\mu m^2 \cdot s^{-1}$.

2) activity of topoisomerase:

The global DNA supercoiling level is held in a homeostatic state by the coordination of two opposing topoisomerases: Gyrase and TopoI. Gyrase responds to positively supercoiled DNA by breaking the DNA double strands and adding two negative supercoils every time, while TopoI only responds to negatively supercoiled DNA and removes one negative supercoil at a time. It is hypothesized that evolution has tuned TopoI to remove the transcription-induced negative supercoils very efficiently since hyper-negative supercoiling is very detrimental to cells, while Gyrase cannot react immediately to the changes brought by transcription, leading to an accumulation of positive supercoils downstream RNAP [5, 27].

Topoisomerase activities can be decomposed into three steps: binding to DNA (with rate k_{bind}), catalysis (with rate k_{cat}), and dissociation (with rate k_{unbind}). For simplicity, we assume that topoisomerase binds indiscriminately throughout DNA, and only the catalysis rate depends on the local supercoiling density. Stracy *et al.* [28] recently characterized the binding kinetics of Gyrase in *E. coli* using single-particle tracking

and diffusion analysis, and they showed that there are about 300 Gyrase molecules stably bound to the DNA at any time, and the average dwell time is about 2 s. Since the *E. coli* genome is about 4 Mb, taking a deterministic approximation, we have $k_{unbind} = 1/(2\text{ s}) = 0.5\ s^{-1}$, and $k_{bind} = (300/4\text{ Mb}\cdot(60\text{ bp/segment}))/(2\text{ s}) = 0.00225\ s^{-1}\cdot segment^{-1}$. Due to the error of deterministic approximation, we further correct k_{bind} to $0.0018\ s^{-1}\cdot segment^{-1}$ by the simulation to match the experimental results. Although the knowledge about TopoI binding kinetics is lacking, we assume that the unbinding rate of TopoI is the same as Gyrase, and we treat the binding rate as a free parameter. The response of catalysis activity to supercoiling density $k_{cat}(\sigma)$ was parameterized from the DNA relaxation assays (**Appendix III**) [5, 29].

3) activities of transcription:

In *E. coli*, a transcribing RNAP is usually a big complex associated with nascent mRNA, ribosomes, and new peptides, and an RNAP is thus, in general, considered as a topological barrier that blocks the diffusion of supercoiling. Suppose an RNAP does not rotate at all, the linking number ahead of and behind it should be conserved. As a result, when an RNAP makes displacement on the DNA (i.e., base stepping), it will push all the turns forward instantaneously and introduce torsional stress. Since the ability to make base stepping is powered by the chemical reaction of NTP and the concentration of NTP is held at constant in most cases, we assume the translocation rate is constant ($60\ bp\cdot s^{-1}$). In addition to displacement, RNAP should also rotate itself to allow certain turns to "diffuse" over it. Although the rotation of RNAP has never been observed in experiments directly, we argue that a slight rotation of RNAP is required to resolve the topological challenge. Since we know very little about the rotation of RNAP, we keep the rotation rate as a free parameter.

To sum up, we assume RNAPs can make two types of motions: displacement and

rotation, where displacement introduces DNA torsional stress, and rotation releases torsional stress. Although it is possible that translocation is accompanied by some conformational changes and slight rotation of RNAP, here we assume that displacement and rotation are uncoupled. Details of the reaction schemes can be found in **Appendix IV**.

2.3 Characterizing the transcriptional response to DNA supercoiling

According to our understanding, DNA supercoiling affects most steps in transcription, including initiation, elongation, and premature dissociation.

1) transcription initiation:

DNA topology has long been considered to affect initiation rate in that supercoiling changes the thermal energy required for the open complex formation [4], which is a crucial step in transcription initiation, and melting is favored over negatively supercoiled DNA. Recently, using *in vitro* single-molecule FISH, Chong *et al.* [5] confirmed that both the T7 RNAP and the *E. coli* RNAP have supercoiling-sensitive initiation rates. These experiments motivated a general theory to quantify the influence of supercoiling on initiation. By adopting a statistical mechanics model, Bohrer *et al.* [30] derived that the initiation rate $k_i(\sigma)$ can be approximated by a linear function with the supercoiling density σ at the promoter. Since this theory quantitatively recapitulated the dynamics of initiation rate observed by Chong *et al.* [5], we utilized it here to describe the supercoiling-sensitive initiation.

$$k_i(\sigma) = \begin{cases} k_{min} & \sigma \geq 0 \\ k_{min} + (k_{max} - k_{min}) \cdot \frac{\sigma}{\sigma_*} & \sigma_* \leq \sigma \leq 0 \\ k_{max} & \sigma \leq \sigma_* \end{cases}$$

We assume that the initiation rate reaches its maximum when DNA is negatively supercoiled to a critical level σ_*, and reaches its minimum when DNA is fully relaxed. σ_* is held at -0.06 for the following simulations.

2) stall during elongation:

Supercoiling is considered to influence the processivity of RNAPs. Experiments suggest that the regulation of supercoiling in elongation can be mediated through the torque-induced stall. During transcription, the change in downstream and upstream DNA supercoiling brought by RNAP translocation will introduce torque to RNAPs. When the torque is beyond some threshold, RNAPs cannot withstand it and will stall instantaneously. After the torque is released, transcription will be resumed. Ma *et al.* [6, 31] has shown that *E. coli* RNAPs can endure torques up to about 10.5 $pN \cdot nm$. We take this value as the stall torque threshold. In our model, the torque that RNAP is subject to is derived from the supercoiling density profile, and the details can be found in **Appendix V**.

Besides torque, elongating RNAP could also stall under extreme supercoiling density. When upstream DNA is hyper-negatively supercoiled, R-loop might form [32], and RNAP could be arrested as a result [33]. Also, supported by some experimental evidence [5], we reasoned that when downstream DNA is hyper-positively supercoiled, the free energy required to melt DNA might be extremely high, and thus RNAP cannot proceed. Here we set ± 0.5 as the supercoiling density threshold that leads to stalled RNAP.

3) premature dissociation:

RNAPs have a chance to prematurely dissociate from DNA under high torsional stress,

as is discovered by Kim *et al.* [7] recently. In the *in vivo* transcription experiment of the lacZ gene, they observed that after the promoter is repressed, the amount of the 5' end and 3' end mRNA product doesn't match, suggesting that already-loaded RNAPs might fall off before they reach the terminator. This phenomenon has been validated in the *in vitro* experiment with RNase, ruling out the possibility of different degradation levels experienced by 5' end and 3' end mRNA in the *in vivo* experiment. Premature dissociation provides a prompt way for cells to stop synthesizing unnecessary products when they shut off the promoter under stress or environmental changes. However, since the phenomenon is discovered only recently, little is known about its kinetics. We simply assume that stalled RNAP will dissociate from DNA with a rate of 0.002 s^{-1}, to match the transcription completion ratio reported in Kim *et al.*[7], which is about 50% for fully-induced LacZ gene after promoter inactivation.

Chapter 3

Results

3.1 Transcription-induced supercoiling can explain the cooperation of RNAPs

To ensure our model can describe the complicated interactions between transcription and supercoiling, we first calibrate it with experimental results and then perform necessary sensitivity analysis. We mainly focus on reproducing Kim *et al.*'s work [7], which is perhaps the most updated and integrative study about single-cell transcription and supercoiling.

Highly expressed genes are often transcribed by multiple RNAPs at the same time, while weakly expressed genes are often transcribed by a single RNAP. A group of co-transcribing RNAPs can display collective behavior, as reported by Kim *et al.* [7], that each of them elongates faster than the case when it moves solo. They termed this phenomenon "cooperation" of RNAPs. A qualitative explanation has been raised: when multiple RNAPs co-transcribe on the same gene, the negative supercoils produced by the leading RNAP and positive supercoils by the trailing RNAP cancel out each other, and the reduced torsional stress makes RNAPs more processive.

To test this explanation quantitatively, we simulated the transcription of lacYZA

(the same systems used in Kim *et al.* [7]) centered in a 9 kb-long open domain, with the chromosome end relaxing to the equilibrium supercoiling density (-0.067) rapidly. For the lacZ gene, when each initiation and termination event happens, a pseudo molecule will be generated separately ("Z5" for initiation and "Z3" for termination). The apparent elongation rate was calculated by dividing the length of the gene over the time interval for the Z3 signal to match the Z5 signal, which is consistent with the method in Kim *et al.* [7]. Note that the apparent elongation rate does not necessarily represent the true elongation rate of RNAP if premature dissociation happens. The co-transcription of a group of RNAPs is simulated using maximum initiation rate k_{max}=0.1 s^{-1}, while solo RNAP transcription is simulated with k_{max}=0.001 s^{-1}. The maximum initiation rate is achieved when $\sigma \leq -0.06$ and minimum initiation of k_{min}=0.001 s^{-1} is achieved when $\sigma \geq 0$. For each setting, 100 replicates were simulated for 500 s. The elongation rate of a group of RNAPs is defined as the average elongation rate of each RNAP in the group; cooperativity is defined as the average apparent elongation rate of a group of RNAPs over that of solo RNAP.

As shown in **Fig. 3-1**, when RNAP rotates slowly (rotation rate <= 5 s^{-1}), a group of RNAPs exhibit cooperative behavior. In these cases, the "cancel-out effects" of supercoils between RNAPs dominate the release of torsional stress: the torque distribution of RNAPs in the group centers at 0 $pN \cdot nm$, while the solo RNAP has a torque distribution centered at around 20 $pN \cdot nm$, which leads to frequent torque-induced stall. However, when RNAP rotates fast, the leaky supercoiling diffusion over RNAP dominates the release of torsional stress: the torque distribution of a group of RNAPs and solo RNAP looks the same, and no cooperation is shown.

In reality, the chromosome end relaxation rate varies, depending on the dynamics of domain barrier formation and the transcription of neighboring genes. To further

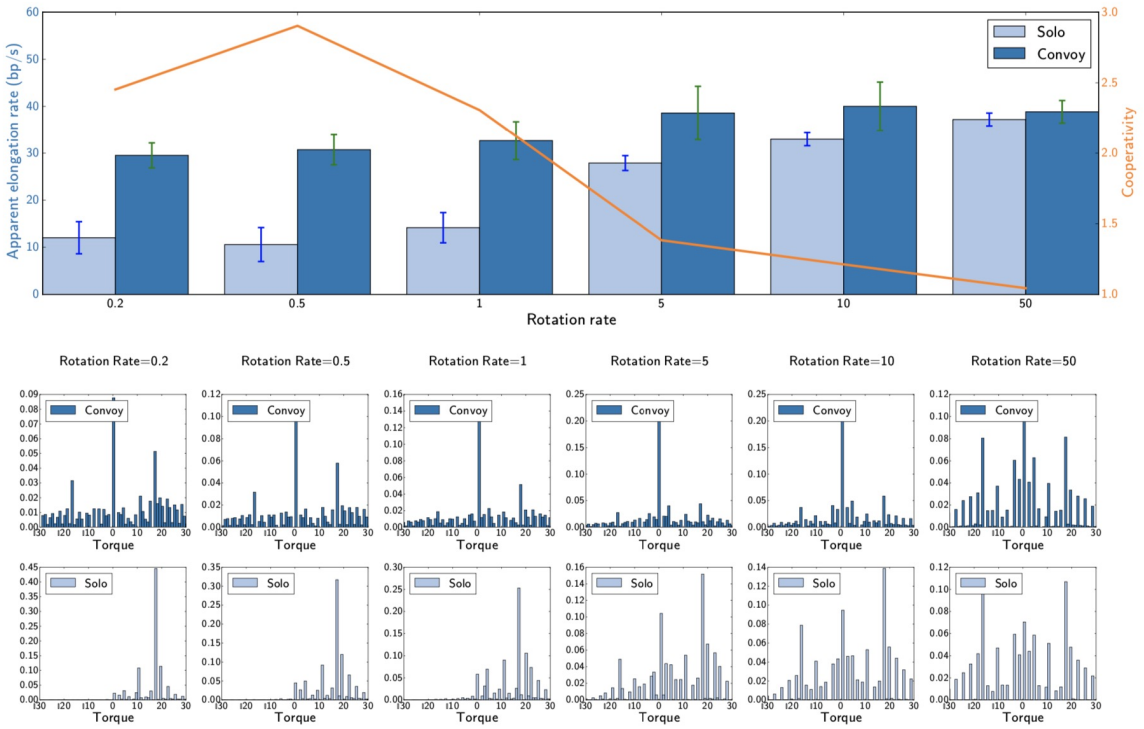

Figure 3-1. Elongation rate, cooperativity, and torque distribution under different RNAP rotation rates (chromosome end relaxation rate = 50). A group of RNAPs is labeled with dark blues bars while solo RNAP is labeled with light blue bars. The upper panel shows the average elongation rate for a group of RNAPs and solo RNAP under different simulation conditions; error bars suggest the standard deviation; the yellow line represents the cooperativity. The lower panel shows the corresponding torque distribution for a group of RNAPs and solo RNAP.

investigate its effect on the cooperative behavior, we performed a sensitivity analysis on both the chromosome end relaxation rate and the RNAP rotation rate. It can be seen from **Fig.** 3-2 that the elongation rate of solo RNAP is insensitive to the chromosome end relaxation rate; rather, it increases as the RNAP rotates faster. However, for a group of RNAPs, when RNAP rotates slowly, the elongation rate differs at different chromosome end relaxation rates. Since the chromosome end relaxation rate only affects the most upstream and most downstream RNAP, we can infer that the behaviors of border RNAPs have a complicated effect on the group dynamics. It is also apparent that cooperativity shows a high dependence on the RNAP rotation

rate, and this feature holds under various chromosome end relaxation rates (even in a closed domain where chromosome end doesn't relax at all).

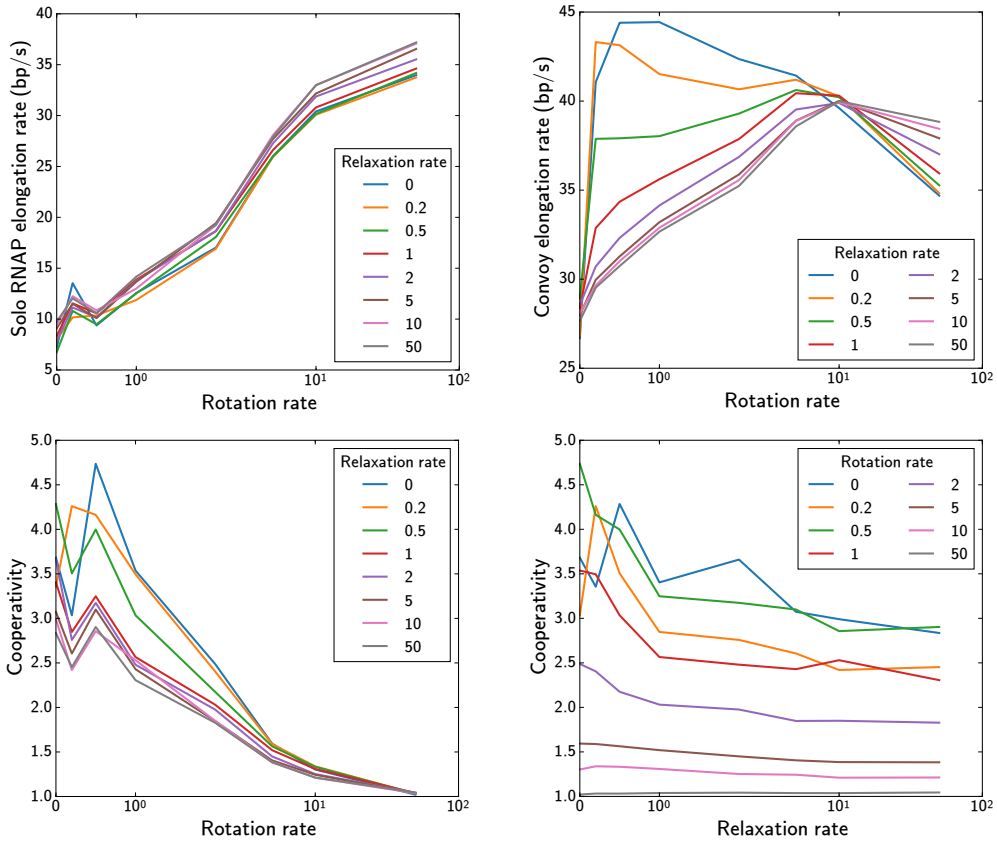

Figure 3-2. Sensitivity of cooperation to RNAP rotation rate and chromosome end relaxation rate. Upper left: solo RNAP elongation rate as a function of RNAP rotation rate. Upper right: the elongation rate of a group of RNAPs as a function of RNAP rotation rate. Lower left: cooperativity as a function of RNAP rotation rate. Lower right: cooperativity as a function of chromosome end rate.

Since TopoI activity also has a significant impact on the supercoiling dynamics, we fixed the RNAP rotation rate at 0.2 s^{-1} and performed a sensitivity analysis on the TopoI activity and the chromosome end relaxation rate. TopoI activity is characterized by TopoI relative binding rate, defined as the TopoI binding rate over the Gyrase binding rate. As we can see from **Fig. 3-3**, both the group of RNAPs and solo RNAP elongate slower when TopoI activity increases. This makes sense since when TopoI activity is too strong, negative supercoils will be removed excessively even before they

cancel out with positive supercoils. When TopoI activity is not too high (equal to or less than fifty times the Gyrase activity), the inter-RNAP supercoiling dynamics are dominated by the "cancel-out effects" of supercoils, and cooperativity is relatively insensitive to TopoI activity. When TopoI activity is relatively high, the inter-RNAP supercoiling dynamics are dominated by the TopoI activity, and thus cooperativity decreases as the TopoI works stronger. This pattern holds under different chromosome end relaxation rates. This result agrees with the observation of Kim *et al.* [7].

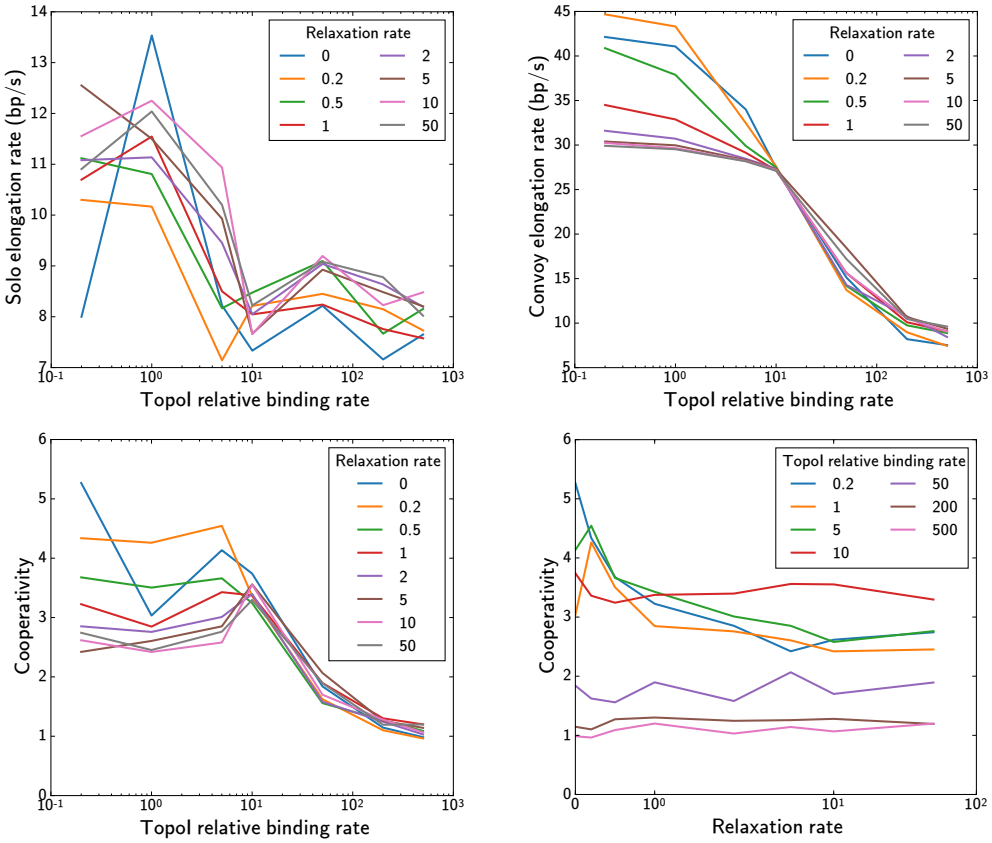

Figure 3-3. Sensitivity of cooperation to TopoI activity and chromosome end relaxation rate. Upper left: solo RNAP elongation rate as a function of TopoI relative binding rate. Upper right: the elongation rate of a group of RNAPs as a function of TopoI relative binding rate. Lower left: cooperativity as a function of TopoI relative binding rate. Lower right: cooperativity as a function of chromosome end rate.

In summary, our model can capture the cooperative behaviors of RNAPs under low RNAP rotation rate, moderate TopoI activity, and various range of chromosome end

relaxation rates. Since the RNAP rotation rate is naturally low (due to sizeable hydrodynamic drag created by nascent transcripts and associated ribosomes)[35], we hypothesize that the cooperative behavior of RNAPs might be a universal feature. For convenience, we fix the RNAP rotation rate as 0.2 s^{-1} and TopoI relative binding rate as 10 for future simulations.

3.2 The regulatory role of initiation on elongation rate

We further investigate the regulatory role of the initiation rate on the elongation rate quantitatively. We simulated genes with different strengths of promoters (with maximum initiation rate k_{max} ranging from 0.001 s^{-1} to 0.1 s^{-1}) and calculated the apparent elongation rate. Average RNAP density, defined as the average number of RNAPs co-transcribing at the same time, was used to characterize the promoter strength. As shown in **Fig. 3-4 (A)**, as RNAP density increases, the elongation rate of each RNAP increases monotonically. This scaling relationship contradicts with the observation in Kim *et al.* [7], where the elongation rate - initiation rate relationship is more "switch-like": as long as RNAPs travel in a group, they exhibit the same elongation rate; increasing the initiation rate will not result in a higher elongation rate.

We also looked at the effects of promoter inactivation on the elongation rate. In addition to cooperation, antagonism also exists in a group of RNAPs, as is reported by Kim *et al.* [7]: after the promoter is inactivated, the group of RNAPs reduces its elongation rate significantly, accompanied by notable premature dissociation. To be specific, in the *in vivo* experiment, they inhibited the *lac* promoter 90 s after induction (about 2700 bp is transcribed) and probed the signal of 5' end mRNA (Z5) and 3' end mRNA (Z3) of the LacZ gene. They observed a significant delay in the Z3 signal

compared to the case without promoter inactivation, indicating the elongation of the already loaded RNAPs is hindered, and a great reduction in the total Z3 signal compared to Z5 signal, indicating premature dissociation of already loaded RNAPs. The antagonistic behavior is hypothesized to be explained by the transcription-induced supercoiling: when the promoter is turned off, there is no continuous supply of RNAPs, and the torsional stress between already loaded RNAPs cannot be alleviated effectively, hampering the processivity of the RNAPs.

To test this idea quantitatively, we simulated the same system and blocked the promoter with an arrested RNAP after 2700 bp is transcribed. As is shown in **Fig. 3-4 (B)**, we cannot reproduce the delay of the Z3 signal compared to the case without premature dissociation. We can, however, if we make the promoter inactivation earlier (for example, inactivating the promoter when 900 bp is transcribed). A possible explanation is that the supercoiling diffusion coefficient in the simulation might be set too low (0.02 $\mu m^2 \cdot s^{-1}$) so that the downstream RNAPs in our model cannot timely "feel" that the promoter is inhibited. However, increasing the supercoiling diffusion coefficient to 0.8 $\mu m^2 \cdot s^{-1}$ does not make antagonistic behavior more significant (data not shown). Another possible reason is that the antagonistic behavior might reinforce as the torsional stress propagate further and further: the downstream RNAP might rotate slower in response to torsional stress [34, 35] since they have longer nascent mRNA and a bigger transcription complex. We will test this hypothesis in the future.

To sum up, our model cannot perfectly reproduce the initiation rate - elongation rate relationship and the negative effects of promoter inactivation on the elongation rate reported by Kim [7], either because of not fully exploring the parameter space, or because of inaccurate or over-simplified model assumptions. The first issue can be addressed by a global fitting using methods like Approximate Bayesian Computation

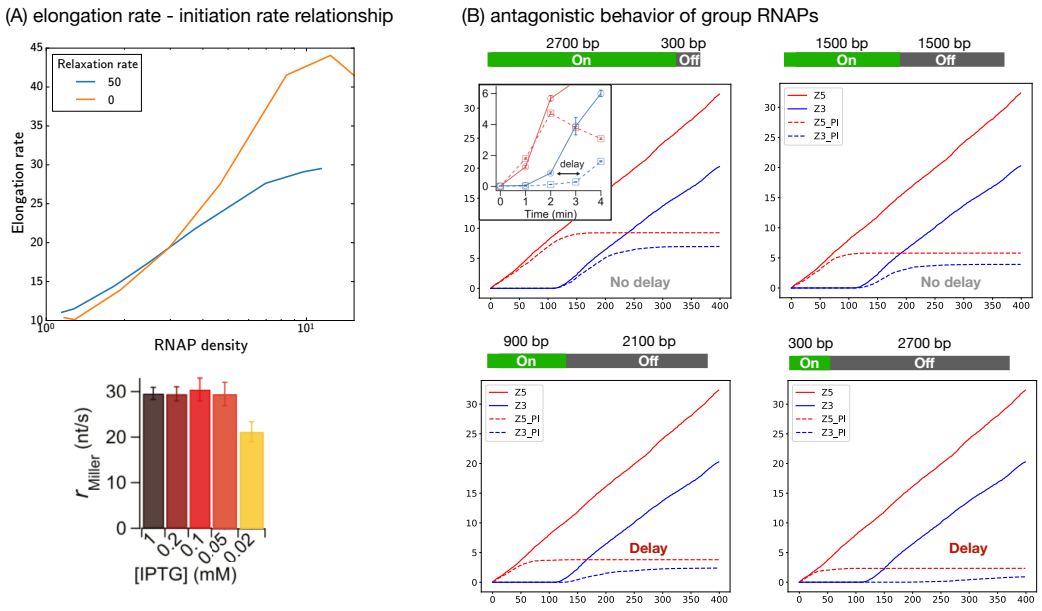

Figure 3-4. Regulation of initiation on elongation rate. (A) Upper panel presents the simulation results. Lower panel shows the measurements from Kim *et al.* [7]. (B) The time evolution of average Z5 and Z3 signals for different simulation settings (x-axis unit: second). Solid lines indicated the simulations without promoter inactivation. Dashed lines indicate the simulations with promoter inactivation. Upper left: promoter is inactivated when the first 2700 bp is transcribed. The inset figure is from Kim *et al.* [7]. Upper right: promoter is inactivated when the first 1500 bp is transcribed. Lower left: promoter is inactivated when the first 900 bp is transcribed. Lower right: promoter is inactivated when the first 300 bp is transcribed.

[36, 37]. The second issue will be reviewed systematically in the discussion section. We argue that, however, our current model is still useful in describing transcription on the mRNA copy number level, since it is able to capture the effects of promoter repression on mRNA production by matching the ratio of premature mRNA. A premature dissociation rate of 0.002 yields a 50% transcription completion ratio after promoter inactivation for strong promoters, agreeing with the experimental data. [7].

3.3 The management of transcriptional noise and the role of topological domains

We further explore the role of promoter strength and topological domains in regulating transcriptional noise. Bacterial transcription exhibits cell-to-cell variability, ensuring the adaptability of cells to various environments [38]. According to Swain *et al.* [39], transcriptional noise can be decomposed into two parts. The noise contributed by the variation in molecule copy numbers in different cells is called extrinsic noise, while the noise coming from the stochastic nature of molecular reactions is called intrinsic noise. Transcriptional noise can be characterized by Fano facor of mRNA copy number, whose lower bound is one, achieved when the mRNA only experiences the stochastic birth-and-death process.

Many genes exhibit greater than one Fano factor in transcription [40, 41], bringing about searches for the sources of transcriptional noise and the physical constraints to it. Previous experiments [42] show that in several constitutive genes, extrinsic noise is enough to explain cell-to-cell variability and the noise level scales with the promoter strength. However, Bohrer *et al.* [30] shows that, if a gene is located in a topological domain, the intrinsic noise brought by the supercoiling-sensitive transcription initiation is also sufficient to explain the cell-to-cell variability. Since this work is performed in a coarse-grained model without spatial resolution, we want to evaluate this idea again using our more accurate spatially-resolved model.

We simulated a 2.4 kb DNA with a 1.2 kb gene locating in the center, with the chromosome end relaxation rate the same as the supercoiling drift rate k_{drift} (which is 50 s^{-1}). As is shown in **Fig. 3-5 (B)**, genes with various promoter strengths produce mRNA with a Fano factor around 1. (Note that the deviation from 1 is a result of the

small number of replicates; more replicates will be simulated in the future.) However, if we put a topological barrier at 0.3 kb and 2.1 kb that loops every 1 min and unloops every 5 min, as we can see from **Fig. 3-5 (C)**, all the genes experience a substantial drop in mean mRNA production, and highly expressed genes (with a k_{max} of 0.04 s^{-1} and 0.10 s^{-1} in this case) exhibit a higher Fano factor.

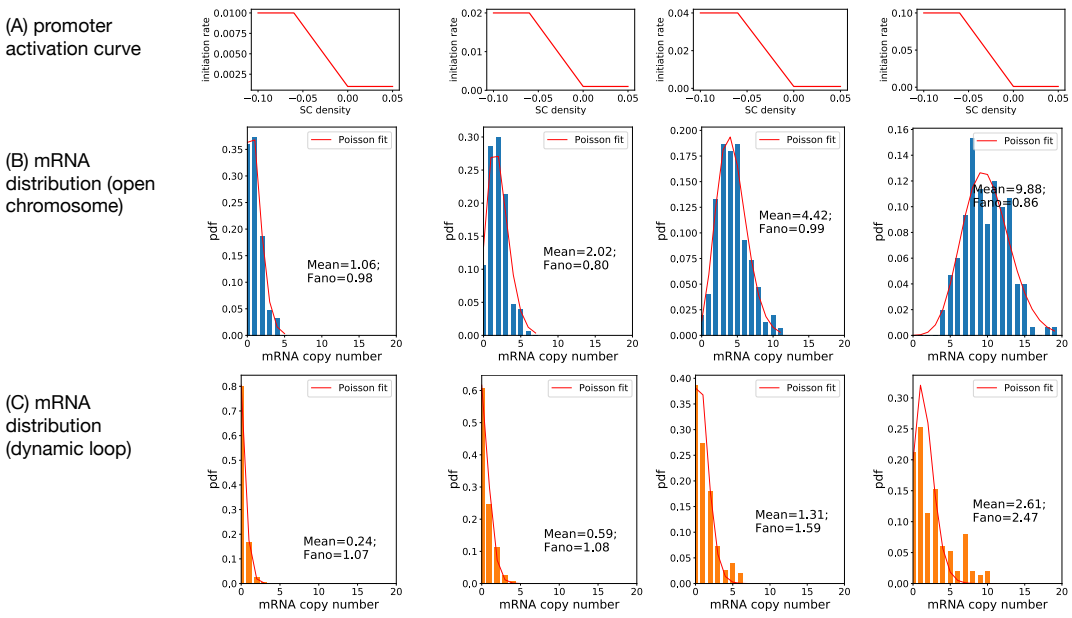

Figure 3-5. Regulatory role of initiation rate on mRNA copy number distribution, in an open chromosome and a dynamic loop respectively. Each column represents the simulation under the same promoter strength setting. For each setting, 150 replicates are simulated and mRNA copy number is counted at t = 2000 s. (A) Promoter activation curve (as a function of supercoiling density at the promoter) for each setting. (B) mRNA copy number distribution simulated in an open chromosome. (C) mRNA copy number distribution simulated in a dynamic loop.

Fig. 3-6 (A) provides a more systematic view of the result. The mRNA of all genes in open chromosome and weakly expressed genes in the dynamic loop are characterized with Fano factor of 1, since the supercoiling level are maintained at equilibrium, either due to chromosome end relaxation or the activities of topoisomerase. However, highly expressed genes are transcribed with larger noise, and the Fano factor

grows with the promoter strength due to the accumulation of torsional stress. As is shown in **Fig. 3-6 (D)(E)**, for strongly expressed genes, the supercoiling density of open chromosome maintains at a moderate level, unaffected by transcription, while the dynamic loop is positively supercoiled due to heavy transcription and inefficient dissipation of excessive supercoils. We also observed that genes in dynamic loop transcribe in a bursty way, characterized by pulses of initiation events. We further analyze the source of burstiness (**Fig. 3-6 (B)(C)**). Bohrer *et al.* [30] shows that the binding and unbinding of Gyrase shared in the topological domain contribute to the transcription bursting. However, Gyrase molecules might bind to DNA more indiscriminately. Here we show that there is no difference in empirical initiation rate when the promoter is bound and when the promoter is unbound by Gyrase, indicating that Gyrase binding is not a source of burstiness. Rather, the empirical initiation rate is quite different in open and closed domains, which means that the loop formation gives rise to the transcription bursting. This makes sense since the torsional stress provides the mechanical "memory" of transcription events, and the release of torsional stress is largely governed by the existence of the topological domain in our model.

Overall, our simulations show that the existence of the topological domain can add intrinsic noise to gene transcription in a promoter strength-dependent way. These results indicate that topological domains act as a transcription regulator, which manages transcription in a global way.

3.4 Information transmission in two-gene system mediated by supercoiling

Last, we explored information transmission between neighboring genes. It has been shown that neighboring genes could regulate the transcription of each other, and the interaction is dependent on the relative orientation of the genes. To investigate

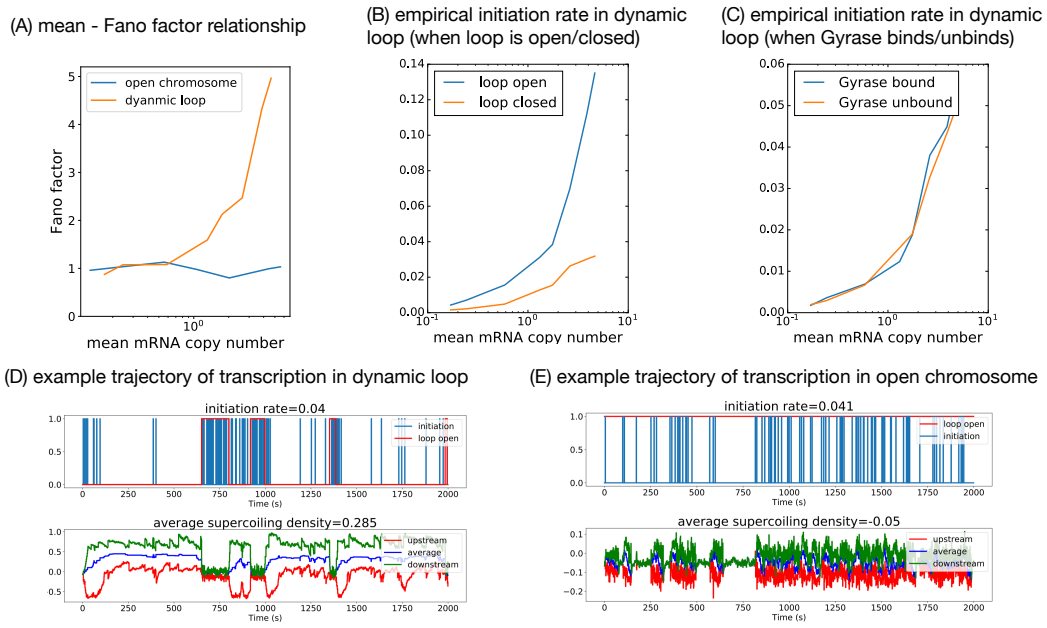

Figure 3-6. Effects of a topological domain on transcription bursting. (A) Fano factor as a function of mean mRNA copy number for genes in an open chromosome and a dynamic loop, respectively. (B) Initiation rate when the loop is open and closed. (C) Initiation rate when the promoter is bound and not bound by Gyrase. (D)(E) Example simulation trajectory. The upper panel represents the initiation events; the lower panel represents the average supercoiling density in different regions (blue: the whole chromosome; red: the region upstream the most left RNAP; blue: the region downstream the rightest RNAP).

this question, we simulated two 1.2 kb genes with the same promoter strength in a linear DNA (**Fig. 3-7 (A)**). Two levels of promoter strength were chosen: k_{max} for the strong promoter is 0.05 s^{-1}, and k_{max} for the weak promoter is 0.01 s^{-1}. The distance of the gene body to chromosome end is 1.2 kb at both sides. Intergenic regions vary in different simulations. We characterize the interaction of two genes by mutual information, which captures the nonlinear correlation between two random variables. Mutual information has been corrected based on sample size.

Consistent with previous predictions [19], when the promoters of two neighboring genes face convergently, the transcription in both genes will generally be inhibited,

compared to the case when they place divergently (**Fig. 3-7 (B)**). This is because positive supercoils generated downstream RNAPs will accumulate in the intergenic region in the convergent genes, and negative supercoils generated upstream RNAP will accumulate in divergent genes. It is also noteworthy that when two genes are very close to each other, even though divergent genes promote the transcription initiation of each other due to excessive negative supercoiling between them, its effects on mRNA production can be overridden by its inhibition on the RNAP elongation rate. When two genes align tandem, the trailing gene will produce positive supercoils downstream and significantly inhibit the transcription initiation of the leading gene, and thus the intergenic region will be net positively supercoiled.

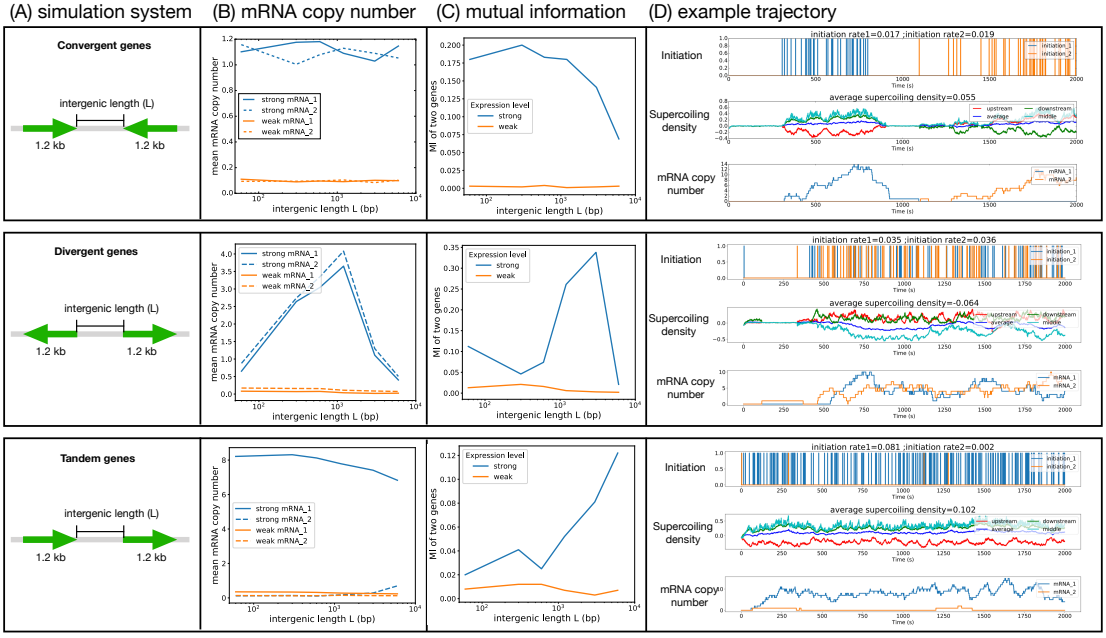

Figure 3-7. Coordination of transcription between neighboring genes. (A) Simulation settings. For each setting, 1200 replicates were simulated and mRNA copy number is counted at t = 2000 s. (B) Mean mRNA copy number of each gene, as a function of intergenic length. (C) Mutual information between mRNA copy number of the two genes, as a function of intergenic length. (D) Example simulation trajectory: the average supercoiling density in the intergenic region is labeled in cyan; the mRNA copy number in each gene are labeled with orange and blue. Other labels are consistent with Fig. 3-6 (D)(E).

Fig. 3-7 (C) also shows some interesting results. For highly expressed genes, information transmission between two genes exhibits an intergenic length-dependent manner. For convergent genes, mutual information decays as the distance between the two genes grows. For divergent genes, mutual information does not decrease monotonically with intergenic length L: it reaches a minimum when L=0.3 kb and reaches a maximum when L=6 kb. A possible explanation is that a negative supercoiled intergenic region confers both advantages and disadvantages to two genes: it boosts the initiation but hampers the elongation. At L= 0.3 kb, the advantages and disadvantages might offset with each other, and thus the interaction of genes is relatively weak; when L=6 kb, the advantageous effects dominate, and the interaction of genes is relatively strong. For tandem genes, mutual information grows as the distance between the two genes grows. A possible reason is that, when two genes locate too closely, the initiation of the leading gene is inhibited almost completely due to the positive supercoils produced by the trailing gene.

To sum up, we briefly analyzed the information transmission of the mRNA production between neighboring genes and presented some interesting predictions which can be tested by experiments. These patterns indicate that neighboring genes also act as a transcription regulator of the target genes, and can be utilized to design better synthetic circuits. For example, for convergent genes, we notice that the transcription of two genes happens in an "either-or" way (**Fig. 3-7 (D)** upper panel), which means that it is possible to use convergent genes to enhance the bistability in synthetic toggle switches [16].

Chapter 4

Discussion

4.1 New insights provided by the model

In this report, we proposed a supercoiling-dependent transcription model in *E. coli*. The biggest difference between our model with previous models is that both transcription initiation and elongation are modeled to be supercoiling-sensitive, while El Houdaigui *et al.* [19] and Ancona *et al.* [20] only modeled the supercoiling-sensitive initiation, and Sevier and Levine [17, 18] only modeled the supercoiling-sensitive elongation. Considering that direct evidence for supercoiling regulation at both initiation and elongation has been shown [5, 6, 31], incorporating both processes allows us to make more accurate quantitative predictions for transcription and allows us to explain phenomena that cannot be explained by previous models. Especially, our model is able to reproduce the orchestration between RNAPs during elongation over long distances [7]. We showed that the cooperation is mediated by supercoiling cancellation and the consequently reduced torque generated by DNA. We further predicted that the cooperation of group RNAPs is a universal behavior if an equilibrium of supercoiling density of -0.06 has been reached in the regions that are adjacent to the gene body. However, genome-scale analysis [43] shows that the supercoiling density in *E. coli* is rather heterogeneous. Future simulations need to be performed to test whether the cooperative behavior still holds when chromosome end supercoiling density changes.

We then utilized this model to explore the role of promoter strength, topological domain, and intergenic context on transcription regulation mediated by supercoiling. Some of the discoveries are consistent with previous models, for example, we showed that topological domains act as a transcription regulator, and the supercoiling dynamics within the domain bring intrinsic noise to genes. However, our model provided different explanations for the underlying mechanism. Previous analysis shows the stochastic binding and unbinding of Gyrase is responsible for the transcription bursting in the topological domain [30]. However, the time scale for Gyrase binding and rebinding in the model is minutes, which contradicts with the recent discoveries shown by single-molecule imaging [28]: the Gyrase binding duration is estimated to be 2-8 s, and there are about 300 Gyrase molecules stably bound in *E. coli* genome at any time, which corresponds to 0.75 molecule at each 10-kb domain (assuming Gyrase binds to DNA indiscriminately). If this is true, there might not be enough time scale separation for supercoiling generation and dissipation, and transcription bursting might not be generated solely by Gyrase activity. Our model showed that it is the formation and dissociation of topological domains that account for the transcription bursts. Also, the half-life for some domains is at the minute scale[44], which provides enough time scale separation for supercoiling generation and dissipation. We did not rule out the role of strong Gyrase binding sites in contributing to transcriptional noise, which might recruit Gyrase with kinetics different from indiscriminate binding. However, strong Gyrase binding sites do not exist in all the topological domains, and we argue that the dynamics of domain formation might be a more prevalent mechanism for transcriptional bursting.

It is worth noting that our model is not perfect. We fail to quantitatively reproduce the delayed transcription termination due to antagonistic behavior of group RNAPs

after promoter inactivation in highly expressed genes. The qualitative reproduction can be achieved, however: in our simulation, the RNAPs do slow down in response to promoter inactivation, and the torsional stress is transmitted towards downstream; the torsional stress is just not transmitted fast enough to make the most downstream RNAP to feel it before it terminated. It is possible that this problem can be resolved by a global fitting in the future. However, if there are some fundamental defects in our model assumptions, a global fitting is not enough; rather, a model selection is needed. We have identified some questionable model assumptions, reviewed in the following section.

4.2 Limitations of the model

The transcription-supercoiling coupling is a rather complicated process, and some assumptions in our model might be overly simplified or imprecise, listed below:

1) supercoiling diffusion and RNAP rotation

Supercoiling diffusion and RNAP rotation are the most important parts of our model since they determine the extent of mechanical memory that can be stored in DNA. In our model, we assume that the supercoiling diffusion coefficient is constant. However, the diffusion of supercoiling is heavily dependent on the force [24], which means that the supercoils near the RNAP might diffuse slower. In addition, the curvature of the chromosome also has impacts on supercoiling diffusion: small natural bends in DNA might introduce a large drag that is sufficient to store torsional stress even in linear DNA that is not topologically constrained [34]. Therefore, the supercoiling density will be more heterogeneous and more dynamic than it is modeled in this report. We also assume that RNAP can rotate itself at a constant rate to release torsional stress. However, RNAPs at different positions are associated with different lengths of nascent mRNA and different amounts of ribosomes, and thus they are likely to

rotate at different rates. Downstream RNAPs might be less sensitive to torque due to higher rotational friction. Incorporating this mechanism might help us reproduce the antagonistic behavior of group RNAPs in the future.

2) RNAP-induced supercoiling

In our model, we update the location of RNAP every 60 bp, and thus transcription-induced supercoiling is generated every 60 bp, which introduces extra (pseudo) noise into our model. Actual RNAP updates its location every single base and generates torsional stress more smoothly. It is possible to model the exact base stepping using our theoretical framework and compare the outcomes of the two. If it turns out that these two models generate similar results, we will use the current 60-bp discretization since it reduces the computational load a lot.

3) interplay between transcription and topological domain

In our model, domain formation is not influenced by supercoiling at the domain boundary. However, evidence shows that transcription activities will change the DNA topology in *Caulobacter crescentus*[45] and *Saccharomyces cerevisiae* [46]. Highly expressed genes will de-compact the chromosome, hinder supercoiling diffusion, and prevent interactions between two chromosome regions. The interaction of transcription and topological domains will add another layer of feedback and complicates information transmission further.

4.3 Future works

In this report, we explore the role of promoter strength, topological domain, and intergenic context in supercoiling-dependent transcription, which regulates gene expression at different scales. The ultimate goal is to understand how the genome organization

and architecture regulates transcription at a whole-genome level. However, investigating this question is beyond our current computational capacity. A better framework to solve this problem is the reaction-diffusion master equation (RDME).

Another major obstacle to quantifying supercoiling-dependent transcription is the lack of knowledge in underlying mechanisms. We still cannot characterize the diffusion of twists on DNA today, and large details of RNAP-DNA interactions are unknown (like the rotational response of RNAP to torsional stress, the mechanism for premature dissociation, etc.). Future single-molecule study on these subjects is required.

References

[1] J. C. Wang, "Interactions between twisted DNAs and enzymes: The effects of superhelical turns," *Journal of Molecular Biology*, vol. 87, no. 4, pp. 797–816, Aug. 1974.

[2] L. F. Liu and J. C. Wang, "Supercoiling of the DNA template during transcription." *Proceedings of the National Academy of Sciences*, vol. 84, no. 20, pp. 7024–7027, Oct. 1987.

[3] H.-Y. Wu, S. Shyy, J. C. Wang, and L. F. Liu, "Transcription generates positively and negatively supercoiled domains in the template," *Cell*, vol. 53, no. 3, pp. 433–440, May 1988.

[4] H. Burns and S. Minchin, "Thermal energy requirement for strand separation during transcription initiation: The effect of supercoiling and extended protein DNA contacts," *Nucleic Acids Research*, vol. 22, no. 19, pp. 3840–3845, 1994.

[5] S. Chong, C. Chen, H. Ge, and X. S. Xie, "Mechanism of Transcriptional Bursting in Bacteria," *Cell*, vol. 158, no. 2, pp. 314–326, Jul. 2014.

[6] J. Ma, L. Bai, and M. D. Wang, "Transcription Under Torsion," *Science*, vol. 340, no. 6140, pp. 1580–1583, Jun. 2013.

[7] S. Kim, B. Beltran, I. Irnov, and C. Jacobs-Wagner, "Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling," *Cell*, vol. 179, no. 1, pp. 106–119.e16, Sep. 2019.

[8] K. Jeong, J. Ahn, and A. B. Khodursky, "Spatial patterns of transcriptional activity in the chromosome of Escherichia coli," *Genome Biology*, vol. 5, no. 11, p. R86, 2004.

[9] P. Sobetzko, "Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes," *Nucleic Acids Research*, vol. 44, no. 4, pp. 1514–1524, Feb. 2016.

[10] J. A. Bryant, L. E. Sellars, S. J. W. Busby, and D. J. Lee, "Chromosome position effects on gene expression in Escherichia coli K-12," *Nucleic Acids Research*, vol. 42, no. 18, pp. 11 383–11 392, Oct. 2014.

[11] S. A. Scholz, R. Diao, M. B. Wolfe, E. M. Fivenson, X. N. Lin, and P. L. Freddolino, "High-Resolution Mapping of the Escherichia coli Chromosome Reveals Positions of High and Low Transcription," *Cell Systems*, vol. 8, no. 3, pp. 212–225.e9, Mar. 2019.

[12] E. Gómez-Díaz and V. G. Corces, "Architectural proteins: Regulators of 3D genome organization in cell fate," *Trends in Cell Biology*, vol. 24, no. 11, pp. 703–711, Nov. 2014.

[13] D. U. Gorkin, D. Leung, and B. Ren, "The 3D Genome in Transcriptional Regulation and Pluripotency," *Cell Stem Cell*, vol. 14, no. 6, pp. 762–775, Jun. 2014.

[14] R. Stadhouders, G. J. Filion, and T. Graf, "Transcription factors and 3D genome conformation in cell-fate decisions," *Nature*, vol. 569, no. 7756, pp. 345–354, May 2019.

[15] L. Postow, "Topological domain structure of the Escherichia coli chromosome," *Genes & Development*, vol. 18, no. 14, pp. 1766–1779, Jul. 2004.

[16] E. Yeung, A. J. Dy, K. B. Martin, A. H. Ng, D. Del Vecchio, J. L. Beck, J. J. Collins, and R. M. Murray, "Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks," *Cell Systems*, vol. 5, no. 1, pp. 11–24.e12, Jul. 2017.

[17] S. A. Sevier and H. Levine, "Mechanical Properties of Transcription," *Physical Review Letters*, vol. 118, no. 26, p. 268101, Jun. 2017.

[18] ——, "Properties of gene expression and chromatin structure with mechanically regulated elongation," *Nucleic Acids Research*, vol. 46, no. 12, pp. 5924–5934, Jul. 2018.

[19] B. El Houdaigui, R. Forquet, T. Hindré, D. Schneider, W. Nasser, S. Reverchon, and S. Meyer, "Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling," *Nucleic Acids Research*, vol. 47, no. 11, pp. 5648–5657, Jun. 2019.

[20] M. Ancona, A. Bentivoglio, C. A. Brackley, G. Gonnella, and D. Marenduzzo, "Transcriptional Bursts in a Nonequilibrium Model for Gene Regulation by Supercoiling," *Biophysical Journal*, vol. 117, no. 2, pp. 369–376, Jul. 2019.

[21] E. Roberts, J. E. Stone, and Z. Luthey-Schulten, "Lattice microbes: High-performance stochastic simulation method for the reaction-diffusion master equation," *Journal of Computational Chemistry*, vol. 34, no. 3, pp. 245–255, Jan. 2013.

[22] M. D. Wang, "Force and Velocity Measured for Single Molecules of RNA Polymerase," *Science*, vol. 282, no. 5390, pp. 902–907, Oct. 1998.

[23] J. Marko and E. Siggia, "Fluctuations and supercoiling of DNA," *Science*, vol. 265, no. 5171, pp. 506–508, Jul. 1994.

[24] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker, "Dynamics of DNA Supercoils," *Science*, vol. 338, no. 6103, pp. 94–97, Oct. 2012.

[25] C. Bécavin, M. Barbi, J.-M. Victor, and A. Lesne, "Transcription within Condensed Chromatin: Steric Hindrance Facilitates Elongation," *Biophysical Journal*, vol. 98, no. 5, pp. 824–833, Mar. 2010.

[26] A. Lesne, J.-M. Victor, E. Bertrand, E. Basyuk, and M. Barbi, "The Role of Supercoiling in the Motor Activity of RNA Polymerases," in *Molecular Motors*, C. Lavelle, Ed. New York, NY: Springer New York, 2018, vol. 1805, pp. 215–232, series Title: Methods in Molecular Biology.

[27] M. Drolet, "Growth inhibition mediated by excess negative supercoiling: The interplay between transcription elongation, R-loop formation and DNA topology," *Molecular Microbiology*, vol. 59, no. 3, pp. 723–730, Feb. 2006.

[28] M. Stracy, A. J. M. Wollman, E. Kaja, J. Gapinski, J.-E. Lee, V. A. Leek, S. J. McKie, L. A. Mitchenall, A. Maxwell, D. J. Sherratt, M. C. Leake, and P. Zawadzki, "Single-molecule imaging of DNA gyrase activity in living *Escherichia coli*," *Nucleic Acids Research*, vol. 47, no. 1, pp. 210–220, Jan. 2019.

[29] R. E. Ashley, A. Dittmore, S. A. McPherson, C. L. Turnbough, K. C. Neuman, and N. Osheroff, "Activities of gyrase and topoisomerase IV on positively supercoiled DNA," *Nucleic Acids Research*, vol. 45, no. 16, pp. 9611–9624, Sep. 2017.

[30] C. H. Bohrer and E. Roberts, "A biophysical model of supercoiling dependent transcription predicts a structural aspect to gene regulation," *BMC Biophysics*, vol. 9, no. 1, p. 2, Dec. 2015.

[31] J. Ma, C. Tan, X. Gao, R. M. Fulbright, J. W. Roberts, and M. D. Wang, "Transcription factor regulation of RNA polymerase's torque generation capacity," *Proceedings of the National Academy of Sciences*, vol. 116, no. 7, pp. 2583–2588, Feb. 2019.

[32] M. Drolet, "The problem of hypernegative supercoiling and r-loop formation in transcription," *Frontiers in Bioscience*, vol. 8, no. 4, pp. d210–221, 2003.

[33] C. Niehrs and B. Luke, "Regulatory R-loops as facilitators of gene expression and genome stability," *Nature Reviews Molecular Cell Biology*, vol. 21, no. 3, pp. 167–178, Mar. 2020.

[34] P. Nelson, "Transport of torsional stress in DNA," *Proceedings of the National Academy of Sciences*, vol. 96, no. 25, pp. 14 342–14 347, Dec. 1999.

[35] Y.-P. Tsao, H.-Y. Wu, and L. F. Liu, "Transcription-driven supercoiling of DNA: Direct biochemical evidence from in vitro studies," *Cell*, vol. 56, no. 1, pp. 111–118, Jan. 1989.

[36] P. Del Moral, A. Doucet, and A. Jasra, "An adaptive sequential Monte Carlo method for approximate Bayesian computation," *Statistics and Computing*, vol. 22, no. 5, pp. 1009–1020, Sep. 2012.

[37] B. M. Turner and T. Van Zandt, "A tutorial on approximate Bayesian computation," *Journal of Mathematical Psychology*, vol. 56, no. 2, pp. 69–85, Apr. 2012.

[38] E. Kussell, "Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments," *Science*, vol. 309, no. 5743, pp. 2075–2078, Sep. 2005.

[39] P. S. Swain, M. B. Elowitz, and E. D. Siggia, "Intrinsic and extrinsic contributions to stochasticity in gene expression," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12 795–12 800, Oct. 2002.

[40] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, "Real-Time Kinetics of Gene Activity in Individual Bacteria," *Cell*, vol. 123, no. 6, pp. 1025–1036, Dec. 2005.

[41] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, "Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells," *Science*, vol. 329, no. 5991, pp. 533–538, Jul. 2010.

[42] D. L. Jones, R. C. Brewster, and R. Phillips, "Promoter architecture dictates cell-to-cell variability in gene expression," *Science*, vol. 346, no. 6216, pp. 1533–1536, Dec. 2014.

[43] A. Lal, A. Dhar, A. Trostel, F. Kouzine, A. S. N. Seshasayee, and S. Adhya, "Genome scale patterns of supercoiling in a bacterial chromosome," *Nature Communications*, vol. 7, no. 1, p. 11055, Apr. 2016.

[44] G. Fulcrand, S. Dages, X. Zhi, P. Chapagain, B. S. Gerstman, D. Dunlap, and F. Leng, "DNA supercoiling, a critical signal regulating the basal expression of the lac operon in Escherichia coli," *Scientific Reports*, vol. 6, no. 1, p. 19243, May 2016.

[45] T. B. Le and M. T. Laub, "Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries," *The EMBO Journal*, vol. 35, no. 14, pp. 1582–1595, Jul. 2016.

[46] Y. J. Achar, M. Adhil, R. Choudhary, N. Gilbert, and M. Foiani, "Negative supercoil at gene boundaries modulates gene topology," *Nature*, vol. 577, no. 7792, pp. 701–705, Jan. 2020.

[47] H. Chen, K. Shiroguchi, H. Ge, and X. S. Xie, "Genome-wide study of mRNA degradation and transcript elongation in E *scherichia coli* ," *Molecular Systems Biology*, vol. 11, no. 1, p. 781, Jan. 2015.

[48] A. Nieß, M. Siemann-Herzberg, and R. Takors, "Protein production in Escherichia coli is guided by the trade-off between intracellular substrate availability and energy cost," *Microbial Cell Factories*, vol. 18, no. 1, p. 8, Dec. 2019.

[49] N. M. Baker, S. Weigand, S. Maar-Mathias, and A. Mondragón, "Solution structures of DNA-bound gyrase," *Nucleic Acids Research*, vol. 39, no. 2, pp. 755–766, Jan. 2011.

[50] B. Krummel and M. J. Chamberlin, "RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes," *Biochemistry*, vol. 28, no. 19, pp. 7829–7842, Sep. 1989.

[51] M. Nöllmann, M. D. Stone, Z. Bryant, J. Gore, N. J. Crisona, S.-C. Hong, S. Mitelheiser, A. Maxwell, C. Bustamante, and N. R. Cozzarelli, "Multiple modes of Escherichia coli DNA gyrase activity revealed by force and torque," *Nature Structural & Molecular Biology*, vol. 14, no. 4, pp. 264–271, Apr. 2007.

[52] J. F. Marko, "Torque and dynamics of linking number relaxation in stretched supercoiled DNA," *Physical Review E*, vol. 76, no. 2, p. 021926, Aug. 2007.

Appendix I

Model species, equations and parameters

Table I-I. Species index

Index	Species
0-(num-1)	DNA
num-(2num-1)	RNAP
2num-(3num-1)	RNAP_stall
3num-(4num-1)	Turn
4num-(5num-1)	Gyrase_unbind
5num-(6num-1)	Gyrase_bind
6num-(7num-1)	TopoI_unbind
7num-(8num-1)	TopoI_bind
8num-(9num-1)	RNAP_tmp
9num-(10num-1)	step
10num-(11num-1)	mRNA
11num-(12num-1)	mRNA_degr
12num-(13num-1)	mRNA_pre
13num-(13num+m-1)	protein
(13num+m)-(13num+2m-1)	S1
(13num+2m)+(13num+3m-1)	S2
13num+3m	loop_state
13num+3m+1	unloop_state
13num+3m+1	loop1
13num+3m+3	loop2
13num+3m+4	unloop2
13num+3m+5	unloop2

Table I-II. Model equations

Description	Reactions	Propensity
	Transcription initiation, elongation and termination	
Transcription initiation	DNA(P) → RNA(P+1)+S1	f_a(Turn(P)) · DNA(P) · DNA(P+1)
RNAP elongation	RNAP(k) → RNAP(k)+step(k)	$k_{elongation}$ · RNAP(k)
RNAP stall	RNAP(k) → RNAP_stall(k)	f_s(Turn(k-1),Turn(k+1)) · RNAP(k)
Stall resumption	RNAP_stall(k) → RNAP(k)	f_r(Turn(k-1),Turn(k+1)) · RNAP_stall(k)
Translocation-1	RNAP(k)+60step(k) → RNAP_tmp(k)	Inf · RNAP(k) · step(k)
Translocation-2	RNAP_tmp(k)+DNA(k+1) → RNAP(k+1)+DNA(k)+mRNA(k)	Inf · RNAP_tmp(k) · DNA(k+1) · DNA(k+2)
Translocation-3	Turn(k) → Turn(k+1)	Inf · Turn(k) · RNAP(k)
RNAP rotation	Turn(k+1) → Turn(k-1)	k_{rot} · RNAP*(k) · Turn(k+1)
RNAP rotation	Turn(k-1) → Turn(k+1)	k_{rot} · RNAP*(k) · Turn(k-1)
Premature termination	RNAP_stall(k) → DNA(k)+mRNA_pre(k)	k_{pre} · RNAP_stall(k)
Termination	RNAP(T) → DNA(T)+mRNA(T)+S2	Inf · RNAP(T)
	mRNA degradation, protein synthesis and protein degradation	
mRNA degradation-1	mRNA(P) → mRNA_degr(P+1)	k_{mdegr} · mRNA(P)
mRNA degradation-2	mRNA(k)+mRNA_degr(k) → mRNA_degr(k+1)	Inf · mRNA(k) · mRNA_degr(k)
mRNA degradation-3	mRNA(T)+mRNA_degr(T) → ∅	Inf · mRNA(T) · mRNA_degr(T)
mRNA degradation-4	mRNA_pre(k)+mRNA_degr(k) → ∅	Inf · mRNA_pre(k) · mRNA_degr(k)
Protein synthesis	mRNA(T) → mRNA(T)+protein	k_{psynt} · mRNA(T)
Protein degradation	protein → ∅	k_{pdegr} · protein
	Supercoiling diffusion, birth and death at chromosome end	
Turn diffusion-1	Turn(k) → Turn(k+1)	k_{drift} · f_d(Turn(k)-Turn(k+1)) · DNA(k+1)
Turn diffusion-2	Turn(k) → Turn(k-1)	k_{drift} · f_d(Turn(k)-Turn(k-1)) · DNA(k-1)
Turn birth	∅ → Turn(end)	k_{birth} · DNA(end)
Turn death	Turn(end) → ∅	k_{death} · Turn(end)
	Topoisomerase binding, catalysis and dissociation	
Gyrase binding	Gyrase_unbind(k)+DNA(k) → Gyrase_bind(k)	k_{gbind} · Gyrase_unbind(k) · DNA(k)
Gyrase catalysis	Turn(k) → ∅	k_{gcat} · f_g(Turn(k)) · Gyrase_bind(k)
Gyrase dissociation	Gyrase_bind(k) → Gyrase_unbind(k)+DNA(k)	k_{gdis} · Gyrase_bind(k)
TopoI binding	TopoI_unbind(k)+DNA(k) → TopoI_bind(k)	k_{tbind} · TopoI_unbind(k) · DNA(k)
TopoI catalysis	∅ → Turn(k)	k_{tcat} · f_g(Turn(k)) · TopoI_bind(k)
TopoI dissociation	TopoI_bind(k) → TopoI_unbind(k)+DNA(k)	k_{tdis} · TopoI_bind(k)
	Topological domain formation and dissociation	

loop formation-1	unloop_state → loop1+loop2+loop_state	$k_{loop} \cdot$ unloop_state
loop formation-2	loop1+DNA(loop1) → ∅	$Inf \cdot$ loop1 $\cdot$ DNA(loop1)
loop formation-3	loop2+DNA(loop2) → ∅	$Inf \cdot$ loop2 $\cdot$ DNA(loop2)
loop dissociation-1	loop_state → unloop1+unloop2+unloop_state	$k_{unloop} \cdot$ unloop_state
loop dissociation-2	unloop1 → DNA(loop1)	$Inf \cdot$ unloop1
loop dissociation-3	unloop2 → DNA(loop2)	$Inf \cdot$ unloop2

Table I-III. Parameter estimation

Notation	Description	Value	Reference
$k_{elongation}$	RNAP elongation rate	60 bp$\cdot s^{-1}$	[5]
k_{rot}	rotation rate of RNAP	0.2 s^{-1}	free parameter
k_{pre}	premature dissociation rate during stall	0.002 s^{-1}	inferred from [7]
k_{mdegr}	mRNA degradation rate	0.0067 s^{-1}	[47]
k_{psynt}	protein synthesis rate	0.03 s^{-1}	[48]
k_{pdegr}	protein degradation rate	0.002 s^{-1}	guess
k_{drift}	supercoiling drift rate	50 s^{-1}	[24]
k_{birth}	supercoiling birth rate at chromosome end	50 s^{-1}	free parameter
k_{death}	supercoiling death rate at chromosome end	1 s^{-1}	free parameter
k_{gbind}	Gyrase binding rate	0.0018 s^{-1}	[28]
k_{gcat}	Gyrase catalytic rate	8.4 s^{-1}	[28]
k_{gdis}	Gyrase dissociation rate	0.4 s^{-1}	[28]
k_{tbind}	TopoI binding rate	0.018 s^{-1}	free parameter
k_{tcat}	TopoI catalytic rate	4.2 s^{-1}	[28]
k_{tdis}	TopoI dissociation rate	1 s^{-1}	[28]
k_{loop}	loop formation rate	0.0167 s^{-1}	free parameter
k_{unloop}	loop dissociation rate	0.0033 s^{-1}	free parameter

Appendix II

Basic modeling strategy

We adopted a stochastic framework since transcription happens in the low copy number regime. Chemical master equations are typically used to describe reactions between species in a probabilistic setting. However, traditional chemical master equations apply to well-stirred systems and cannot capture the spatial heterogeneity of molecules. To overcome this, we partitioned the DNA into segments of a fixed size. Each segment has several attributes (for example, the availability of DNA, the occupation of normal RNAP, the occupation of stalled RNAP and the number of turns), and different attributes at each segment is represented by different species. Evolution of attributed at different segments can be represented by different reactions. Attributes can be either binary or integral.

In the case mentioned above, the availability of DNA, the occupation of normal RNAP and the occupation of stalled RNAP are binary attributes, and they are mutually exclusive. If $RNAP(k) = 1$, both $DNA(k)$ and $RNAP\_stalll(k)$ will be 0, since the DNA at k-th segment has been occupied by RNA, and the RNAP at k-th segment can only take one state, either normal or stalled. Similarly, if $DNA(k) = 1$, both $RNAP(k)$ and $RNAP\_stalll(k)$ will be 0, since $DNA(k) = 1$ suggests that no proteins occupy the k-th segment. The number of turns is an integral attributed: if $Turn(1) = 5$, it means that 5 turns wraps around the first DNA segment, and we can calculate the supercoiling density accordingly.

This framework shown above is very flexible and can be applied to any spatially resolved dynamical system. The drawback is that as we increase the number of the segments we simulate, the number of species and reactions grow proportionally, increasing the simulation time. To balance the accuracy and computational load, we partitioned the DNA into 60-bp segments. We choose 60-bp as the unit of positions since it is at the same magnitude of the size of Gyrase binding sites (137 bp) [49] and RNAP footprints (32 bp) [50].

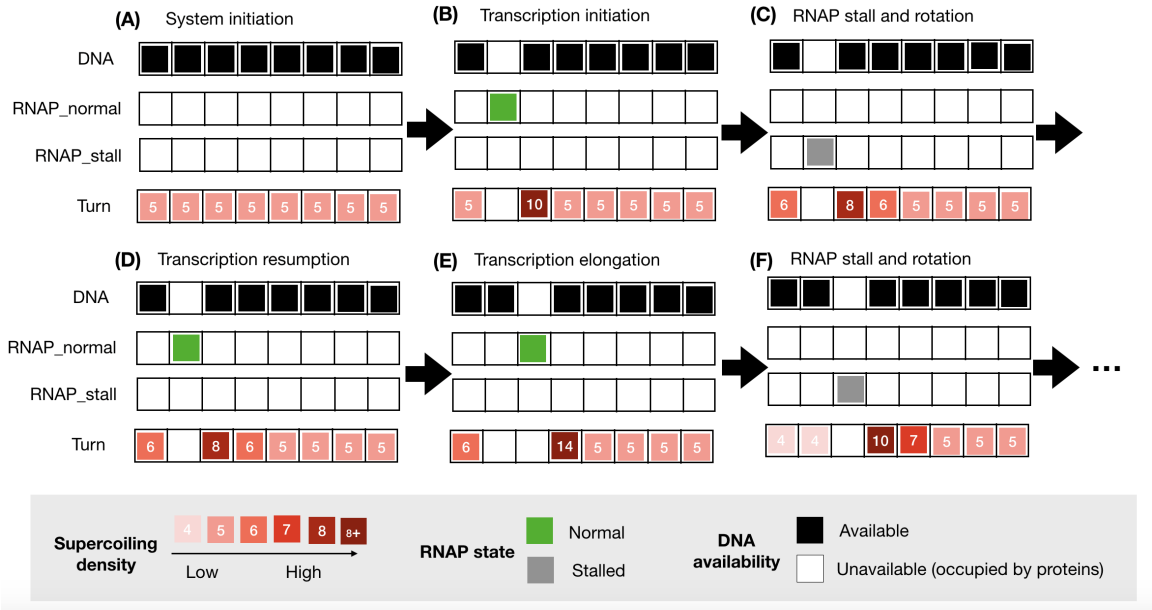

Figure II-1. A toy model that shows how spatially resolved model works. (A) The system is initialized with all DNA available and evenly distributed turns. (B) Once transcription initiates, the 2nd DNA segment is occupied and the turns are pushed forward. (C) Due to the high torsional stress, RNAP switches to the stalled state and allows the turns to diffuse over it. (D) Due to the reduced torsional stress, the RNAP switches back to the normal state and is ready for translocation. (E) The RNAP translocates to the 3rd segment and pushes turns forward. (F) The RNAP switches to the stalled state and allows the turns to diffuse over it.

Appendix III

Characterizing the response of topoisomerase to supercoiling

Both Gyrase and TopoI are supercoiling-sensitive. It has been shown that Gyrase removes positive supercoils faster than it introduces negative supercoils in relaxed DNA [29]. Moreover, TopoI could only work when DNA is negatively supercoiled [5]. In this work, since we only care about the effects of topoisomerase on the supercoiling density, the detailed kinetics were not taken into consideration (like the multiple modes of Gyrase [51]). Instead, we used phenomenological models to characterize the overall activity of topoisomerase in response to DNA supercoiling density. We assumed that the unbinding rate and catalytic rate are invariant, and the binding rate of topoisomerase changes when DNA supercoiling density σ changes.

The workflow is as follows: we first convert the image of ensemble DNA relaxation assay to the intensity of bands using ImageJ. The bands on the DNA relaxation assay show how DNA supercoiling density changes overtime after a specific kind of topoisomerase is added into a supercoiled DNA. We assume that the intensity of bands has a linear relationship with the average supercoiling density. We then infer the supercoiling density at each time point from the intensity data, which is further fitted with an ODE that characterizes the activity of topoisomerase.

For Gyrase, we assumed that its binding rate follows a sigmoidal function of σ. From (**Fig. III-1**), we can see that the sigmoidal curve fits the experimental data from Ashley *et al.* [29] well. Since the amplitude (k_1) of this sigmoidal curve is dependent on the specific concentration of Gyrase used in the experiment, we only used k_2 and k_3 to represent the shape of the curve. In our study, k_1 is inferred from the *in vivo* experiments[28], which has been discussed in Methods. For TopoI, we assumed that its binding rate follows a stepwise function of σ: when DNA is negatively supercoiled, the binding rate is a constant; when DNA is relaxed or positively supercoiled, the binding rate is zero. It fits the experimental data from Chong *et al.* [5] well (**Fig. III-2**).

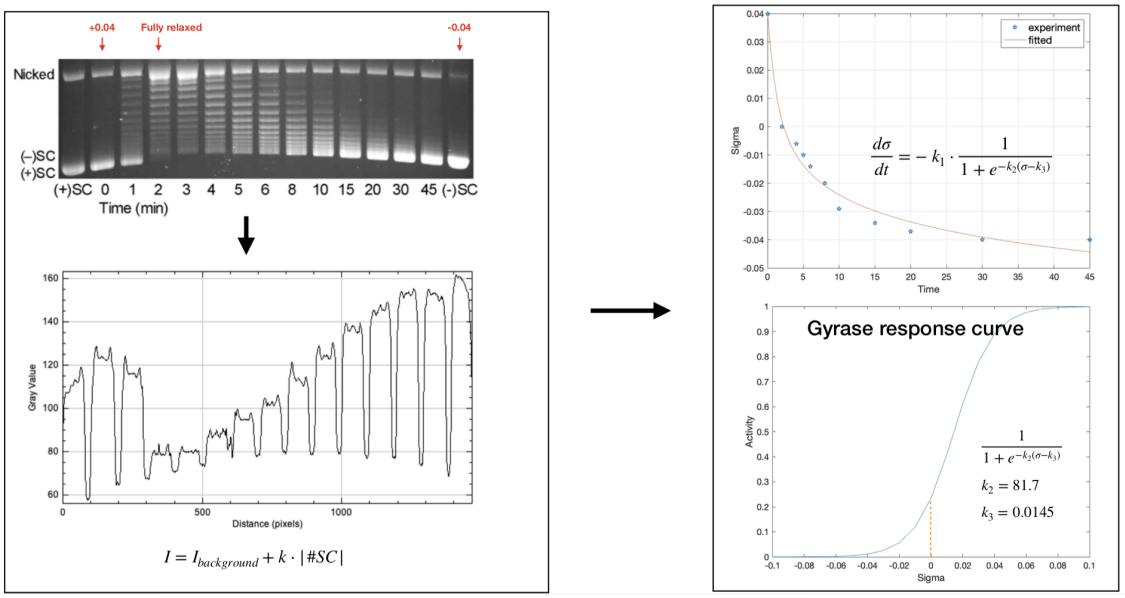

Figure III-1. Calibration of Gyrase's response to supercoiling. Left panel: conversion from gel image to band intensity. Right panel: fitting the supercoiling density data with a sigmoidal curve that represents the Gyrase activity.

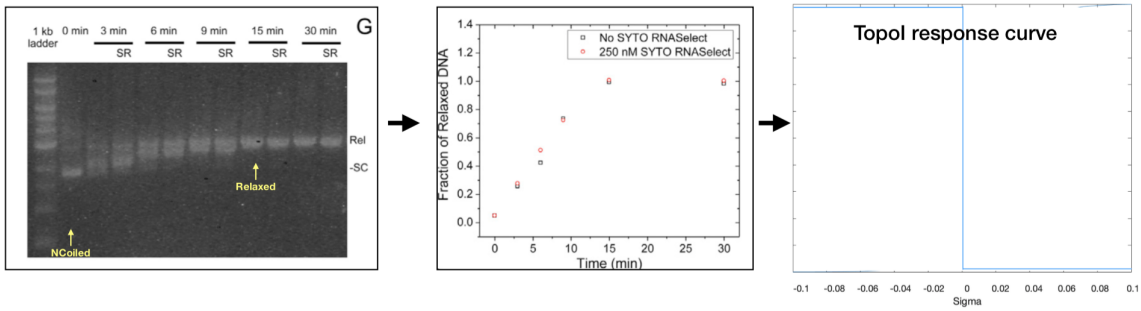

Figure III-2. Calibration of TopoI's response to supercoiling. Left panel: conversion from gel image to band intensity. Right panel: fitting the supercoiling density data with a stepwise function that represents the TopoI activity. (Before relaxation, the changes in supercoiling density over time is linear, indicating that the TopoI activity is constant.)

Appendix IV

Characterization of RNAP motions

1) rationale for RNAP rotation

We argue that during transcription, an RNAP must allow certain twists or writhes to "diffuse" over it, which is through the rotation of RNAP. This argument can be deduced from the *in vitro* observation made by Chong *et al.* [5] that transcription initiation rates remained constant in the absence of topoisomerase, where multiple biotins fixed the circular DNA template on the surface. In this case, if RNAPs cannot rotate, as RNAPs approach the biotin tags (topological barriers), an extreme level of positive supercoiling will build upstream to the extent that the transcription cannot proceed, and all the RNAPs will stall out, and no continuous initiation will happen, which contradicts the experimental observation. The experiments suggest that apart from topoisomerase, there must be other mechanisms to release the torsional stress, and we think it can be either achieved by RNAP rotation or RNAP premature dissociation. The rotation of RNAP can be modeled equivalently to the diffusion of supercoiling over a topological barrier:

$$\text{RNAP(k) + Turn(k-1)} \rightarrow \text{RNAP(k) + Turn(k+1)}$$
$$\text{RNAP(k) + Turn(k+1)} \rightarrow \text{RNAP(k) + Turn(k-1)}$$

2) equations for RNAP displacement and rotation

To avoid collisions between RNAPs, we only allow displacement to happen when the forward two segments (namely, DNA(k+1) and DNA(k+2)) are not occupied by other RNAPs. Thus, the displacement of RNAP from segment k to segment k+1 can be modeled in this way:

$$\text{RNAP(k) + DNA(k+1) + DNA(k+2)} \rightarrow \text{RNAP(k+1) + DNA(k) + DNA(k+2)}$$
$$\text{Turn(k)} \rightarrow \text{Turn(k+1)}$$

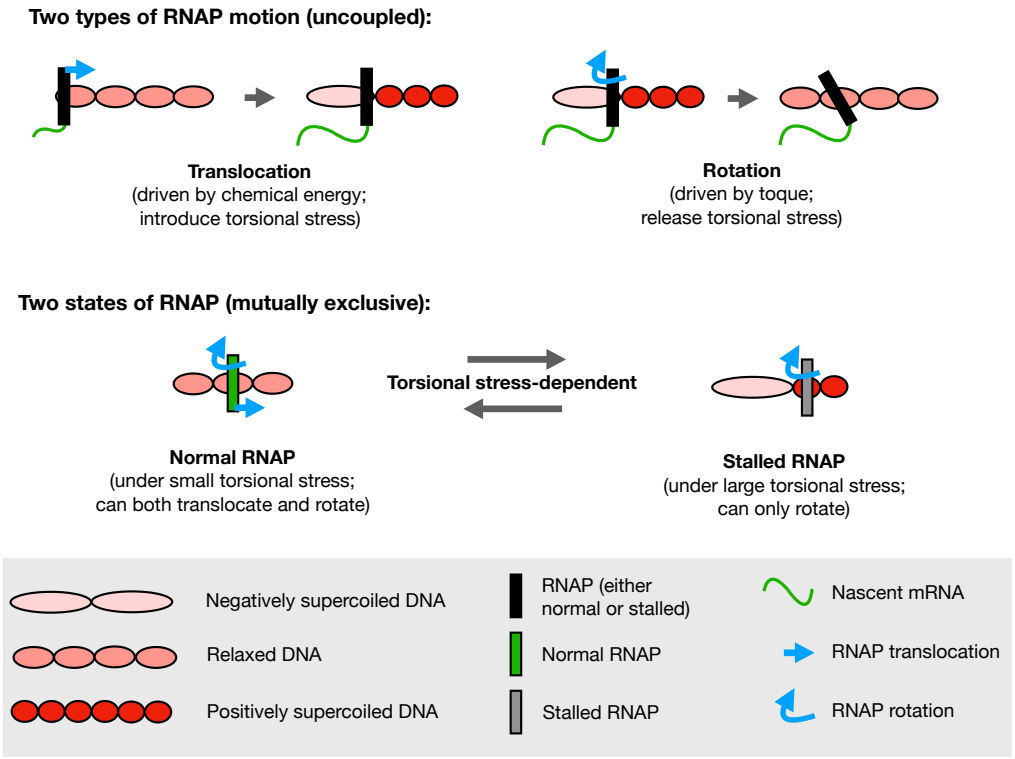

Figure IV-1. Assumption about motions and states of RNAP.

Appendix V

Calculating supercoiling density and torques

The natural linking number lk_0 for a 60-bp DNA is about 5.7 (considering relaxed DNA makes a turn every 10.5 bp). Suppose the actual linking number on a specific segment is x, then the average supercoiling density on this DNA segment could be calculated by $\frac{x-5.7}{5.7}$. In our model, we use "Turn(k)" to represent the linking number of segment k. Since the quantities in chemical master equations should be integers rather than real numbers, we define Turn(k) as $10x$ and define the new lk_0 as 57. The supercoiling density is $\frac{10x-57}{57} = \frac{x-5.7}{5.7}$. To initialize a negatively supercoiled DNA with $\sigma = -0.05$, we set Turn(k) = 54 for all k.

Marko [52] proposed a quantitative framework for calculating the torque generated by a DNA held with a constant force of F:

$$\tau = \begin{cases} \frac{c_s}{\omega_0} \cdot \sigma & |\sigma| \leq |\sigma_s| \\ [2pg/(1-p/c_s)]^{\frac{1}{2}}/\omega_0 & \sigma_s \leq \sigma \leq |\sigma_p| \\ \frac{p}{\omega_0} \cdot \sigma & |\sigma| \geq |\sigma_p| \end{cases}$$

The coefficients c_s, p, g and critical values σ_s, σ_p can be calculated from mechanical properties of DNA:

$$p = k_B T P {\omega_0}^2$$

$$c = k_B T C {\omega_0}^2$$

$$c_s = c(1 - \frac{C}{4L_p}\sqrt{\frac{k_B T}{L_p F}})$$

$$g = F - \sqrt{\frac{k_B T F}{L_p}}$$

$$\sigma_s = \frac{1}{c_s}(\frac{2pg}{1-p/c_s})^{\frac{1}{2}}$$

$$\sigma_p = \frac{1}{p}\left(\frac{2pg}{1-p/c_s}\right)^{\frac{1}{2}}$$

Here we choose temperature (T) as 298 K, twist persistence length (C) as 95 nm, stiffness in supercoiled DNA (P) as 24 nm, DNA bending persistence length (L_p) as 50 nm, the contour-length rate of rotation of the relaxed double helix (ω_0) as $2\pi/(3.6nm) = 1.76\ nm^{-1}$, and force ($F$) as 0.15 pN.

With this framework, given the number of turns immediately downstream and upstream of RNAP, we can convert supercoiling density to torque on each side. For RNAP, the overall torque it is subject to is calculated from the downstream torque minus the upstream torque.

Curriculum Vitae **Yuncong Geng** **May, 2020**

2910 Glen Ave Apt B
Baltimore, Maryland 21215 USA
(+1) 410-900-5827
ygeng7@jhu.edu

Education

M.S.E in Biomedical Engineering
Johns Hopkins University, Baltimore *(2018–present)*

B.Eng. in Biomedical Engineering
Southeast University, Nanjing, China *(2013–2017)*

Research Experience

Stochastic modeling of supercoiling-dependent transcription in E. coli (Advisor: Elijah Roberts)
Johns Hopkins University, Baltimore *(April 2019–present)*

- Developed a chemical master equation-based, spatially resolved and supercoiling-dependent model for E.coli transcription
- Validated the model by reproducing results from main experimental studies on supercoiling-dependent transcription
- Explored the collective behaviors of RNAPs
- Characterized the effects of topological domain on transcriptional noise
- Investigated the role of adjacent gene on regulating the transcription of target gene, and how it is affected by the orientation of the genes

Sample selection from large RNA-seq datasets (Advisor: Patrick Cahan)
Johns Hopkins University, Baltimore *(Sept 2018–Dec 2018)*

- Designed a graph-based method to select samples with desired developmental distance from a cell-type ontology tree
- Implemented a hierarchical clustering-based sample selection method
- Tested the methods on datasets with 173 cell types represented by 573 samples

Genomic analysis of the evolutionary origin of multiple sex chromosomes in monotreme (Supervisor: Qiye Li)
BGI, Shenzhen, China *(Jun 2016–Jun 2018)*

- Evaluated the performance of genome assembly algorithm for third-generation sequencing data, and developed a method to identify artificial segment duplication resulting from assembly error of second-generation sequencing data
- Participated in genome assembly, annotation, expression analysis and phylogenetic analysis of platypus and shorted-beaked echidna
- Identified successive recombination suppression events and the associated genes on sex chromosomes, and predicted potential sex determining locus

Development of test strip for early vascular inflammation marker – homocysteine (Hcy) in blood (Advisor: Yu Zhang)
Southeast University, Nanjing, China *(2015–2016)*

- Raised a proposal to detect blood Hcy, by first converting Hcy to S-Adenosyl-L-homocysteine (SAH) which is larger and easier to detect, and then capturing SAH using competitive immunochromatographic assay

- Optimized the synthesis conditions of competitive BSA-coated antigen and characterized the specificity and sensitivity of anti-SAH antibodies

Additional Skills

Computing and Other Skills

- Proficient in Python, R, Unix Shell, MATLAB and Perl
- Experiences with C++ and SBML

Molecular Techniques

- Familiar with bio-analytical technologies: UV/Vis spectroscopy, electrophoresis, ELISA and flow cytometry
- Experiences with plasmid-based genetic engineering on E. coli, synthesis and characterization of nanoparticles

Teaching Experience

Teaching Assistant for Molecular & Cellular Systems Biology (AS.250.313)
Johns Hopkins University, Baltimore *(Sept 2019-Dec 2019)*

- Responsible for grading, holding office hours and answering emails
- Topics: dynamic modeling of biochemical pathways, steady-state analysis of metabolic networks, inference of gene regulatory networks, and systems biology approaches to studying signal transduction